

12-12-2019

Power Quality Disturbance Recognition Base on Maximum Variance Unfolding

Linlin Che

College of Electrical and Electronic Engineering, North China Electric Power University, Baoding 071003, China;

Yinghui Kong

College of Electrical and Electronic Engineering, North China Electric Power University, Baoding 071003, China;

Zhixiong Chen

College of Electrical and Electronic Engineering, North China Electric Power University, Baoding 071003, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research](#), [Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Power Quality Disturbance Recognition Base on Maximum Variance Unfolding

Abstract

Abstract: In view of the complexity of power quality disturbance (PQD), a recognition method of PQD feature extraction method based on maximum variance unfolding (MVU) is proposed and PQD recognition is completed by combining classifier algorithm in this paper. *The wavelet energy features of PQD are extracted by wavelet transform to construct the original feature set. For the reason that non-convex quadratic programming is transformed into convex semi-definite optimization problem by kernel function, the MVU method is applied to compress the sample features into 3-dimension features which keep low dimensional structure in high dimensional space and make a rough recognition of PQD. General classifiers are used to verify the method.* The experimental results show that this MVU method reduces the number of eigenvectors and improves the recognition accuracy of the PQD. It is promising in engineering.

Keywords

power quality disturbance, manifold learning, maximum variance unfolding, features extraction, disturbance recognition

Recommended Citation

Che Linlin, Kong Yinghui, Chen Zhixiong. Power Quality Disturbance Recognition Base on Maximum Variance Unfolding[J]. Journal of System Simulation, 2019, 31(8): 1702-1710.

基于最大方差展开法的电能质量扰动识别

车麟麟, 孔英会, 陈智雄

(华北电力大学 电气与工程学院, 河北 保定 071003)

摘要: 针对电能质量扰动(power quality disturbance, PQD)的复杂性, 提出了一种基于最大方差展开(maximum variance unfolding, MVU)非线性流行学习的 PQD 特征提取方法, 结合分类器算法完成了对 PQD 识别。对 PQD 信号进行小波分解得到信号的小波能量作为原始特征集;通过 MVU 算法对原始特征集进行压缩,由于在算法中引入核函数将非凸二次规划转化为凸半正定最优化问题,从而得到信息量更集中且很好保持训练数据分布边界的低维 PQD 特征;结合分类器算法完成 PQD 识别。实验结果表明, MVU 算法约简后得到 PQD 特征向量,不仅有效降低了特征向量个数,而且对 PQD 的识别准确率高,有一定的工程应用前景。

关键词: 电能质量扰动; 流行学习; 最大方差展开; 特征提取; 扰动识别

中图分类号: TM711

文献标识码: A

文章编号: 1004-731X (2019) 08-1702-09

DOI: 10.16182/j.issn1004731x.joss.17-0295

Power Quality Disturbance Recognition Base on Maximum Variance Unfolding

Che Linlin, Kong Yinghui, Chen Zhixiong

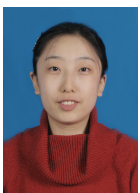
(College of Electrical and Electronic Engineering, North China Electric Power University, Baoding 071003, China)

Abstract: In view of the complexity of power quality disturbance (PQD), a recognition method of PQD feature extraction method based on maximum variance unfolding (MVU) is proposed and PQD recognition is completed by combining classifier algorithm in this paper. The wavelet energy features of PQD are extracted by wavelet transform to construct the original feature set. For the reason that non-convex quadratic programming is transformed into convex semi-definite optimization problem by kernel function, the MVU method is applied to compress the sample features into 3-dimension features which keep low dimensional structure in high dimensional space and make a rough recognition of PQD. General classifiers are used to verify the method. The experimental results show that this MVU method reduces the number of eigenvectors and improves the recognition accuracy of the PQD. It is promising in engineering.

Keywords: power quality disturbance; manifold learning; maximum variance unfolding; features extraction; disturbance recognition

引言

随着大量的非线性负载和电力电子设备接



收稿日期: 2017-06-28 修回日期: 2017-09-18;
基金项目: 国家自然科学基金(61771195, 61601182),
中央高校科研业务费资助(2015MS100), 河北省自然
科学基金 (F2017502059);
作者简介: 车麟麟(1981-), 女, 河北保定, 硕士,
高工, 研究方向为电能质量分析等。

入电网中, 为了保障电能质量满足用户的性能要求, 有必要对输电和配电系统中的电能质量进行有效检测和分析^[1-3]。其中, 电能质量扰动(Power Quality Disturbance, PQD)信号是分析电能质量的重要参数, PQD 识别技术已成为电能质量分析领域的一个重要研究方向。PQD 识别的目的是从海量的电能质量数据中将 PQD 快速准确地定位和识

别。PQD 识别过程包括特征提取和模式识别两部分。而 PQD 特征提取则是 PQD 识别的关键所在^[4], 好的 PQD 特征能有效地提高识别准确度和减少计算复杂度。

PQD 特征提取是通过映射变换提取到能反映扰动信号波形特征的特征量。目前, 常用的特征提取方法包括: 傅里叶变换^[5]、小波变换^[6]、希尔伯特-黄变换等。其中, 傅里叶变换着重反映分析信号的整体信息, 但是由于忽略了信号的局部特性, 且对于非平稳信号不具备时间局部性, 因此不满足时频分析要求。虽然小波变换已被广泛应用于 PQD 特征提取, 该变换从各层小波分解系数中提取特征向量, 适用于平稳和非平稳信号进行分析, 可获得较好的效果, 但是小波变换的算法复杂度较高, 仍有进一步被压缩简约的空间。因此, 本文考虑将数据降维技术应用到 PQD 特征向量简约的过程中。

数据降维技术分为线性降维和非线性降维, 非线性降维又分为基于核方法和基于几何结构的非线性降维。几何结构的非线性降维, 即流行学习可在有效减少数据规模的同时保留高维空间的非线性流形结构, 实现维数约简。并且流形学习是在没有任何先验知识的条件下, 根据高维观测数据集发现未知映射, 并找到与高维观测数据集对应的低维流形表示。它是从观测到的现象中寻找事物的本质, 找到产生数据的内在规律的一种方法。目前主要的流形学习算法^[7]有如下几个: 等距映射算法(Isometric Mapping, ISOMAP)、局部线性嵌入算法(Locally Linear Embedding, LLE)、局部切空间排列算法(Local Tangent Space Alignment, LTSA)、局部保持映射算法(Locality Preserving Projection, LPP)和最大方差展开算法(Maximum Variance Unfolding, MVU)。其中, MVU 算法是一种基于流形局部等距概念的流形学习算法^[8], 当近邻点间的欧式距离保持不变时, 可通过平移旋转等变换, 在低维空间中展开高维数据流形。MVU 算法能有效的学习出隐含在高维数据集中的低维流形, 并且

对于非凸数据集有较好的降维效果^[9]。MVU 降维后的低维数据较好地保持了训练数据的分布边界^[10]。

在 PQD 识别领域对于基于几何结构的非线性降维鲜有研究, 因此本文对该技术在 PQD 识别中的应用进行了全新的探索和尝试。结合小波变换出现的 PQD 识别问题, 本文提出了一种基于 MVU 方法的 PQD 特征提取方法。考虑扰动参数随机性和噪声影响, 论文根据常见的 PQD 模型产生信号, 在提取信号小波能量向量的基础上通过 MVU 进行特征向量降维, 得到的低维 PQD 特征向量很好地保持了原数据的分布边界, 且信息量更集中。最后结合经典分类器进行 PQD 识别。仿真和实测实验结果表明该方法提取到的特征向量能很好地实现对 PQD 信号的识别。由于 MVU 算法缩减了特征向量个数并满足 k -最近邻点间局部距离不变的约束, 在选取核函数时充分考虑了高维数据集的分布特性, 因此减轻了后续 PQD 的分类压力, 减少了分类运算时间, 提高了 PQD 识别准确率。

1 PQD 信号模型

由于电能质量信号复杂且种类较多, 因此国内外文献大都采用仿真法得到数据样本进行相关分析。本文参考文献[11]建立标准信号与 6 种扰动信号的仿真模型, 见表 1。

表 1 PQD 信号模型
Tab. 1 Signal model of PQD

类型	参数	公式
正弦		$v(t)=\sin(\omega t)$
凸起	$0.1 \leq \alpha \leq 0.8,$ $T \leq t_2 - t_1 \leq 9T$	$v(t)=A\{1+\alpha[u(t_2)-u(t_1)]\}\sin\omega t$
凹陷	$0.1 \leq \alpha \leq 0.8,$ $T \leq t_2 - t_1 \leq 9T$	$v(t)=A\{1-\alpha[u(t_2)-u(t_1)]\}\sin\omega t$
间断	$0.9 \leq \alpha \leq 1,$ $T \leq t_2 - t_1 \leq 9T$	$v(t)=A\{1-\alpha[u(t_2)-u(t_1)]\}\sin\omega t$
谐波	$0.05 \leq \alpha_3, \alpha_5, \alpha_7 \leq 0.15,$ $\sum \alpha_i^2 = 1$	$v(t)=A[\alpha_1\sin\omega t + \alpha_3\sin 3\omega t + \alpha_5\sin 5\omega t + \alpha_7\sin 7\omega t]$
脉冲		$v(t)=\sin\omega t + a\delta(t-t_0)$
振荡	$0 \leq t_2 - t_1 \leq 2T$	$v(t)=\sin\omega t + ae^{-c(t-t_1)}\sin b\omega t[u(t_2)-u(t_1)]$

表 1 提供的数学公式模型包括了 6 种常见的 PQD 信号(电压凸起、电压凹陷、电压间断、谐波、电压暂态, 其中电压暂态包括脉冲暂态和振荡暂态)。其中: $u(t)$ 为阶跃函数; α 为幅度; t_1 和 t_2 分别为扰动开始时刻和结束时刻; T 为信号周波; ω 为基波角频率, 基波频率设定为 50 Hz。其输入信号的分析时间长度取为 10 个周波即 0.2 s。其波形图如图 1 所示。

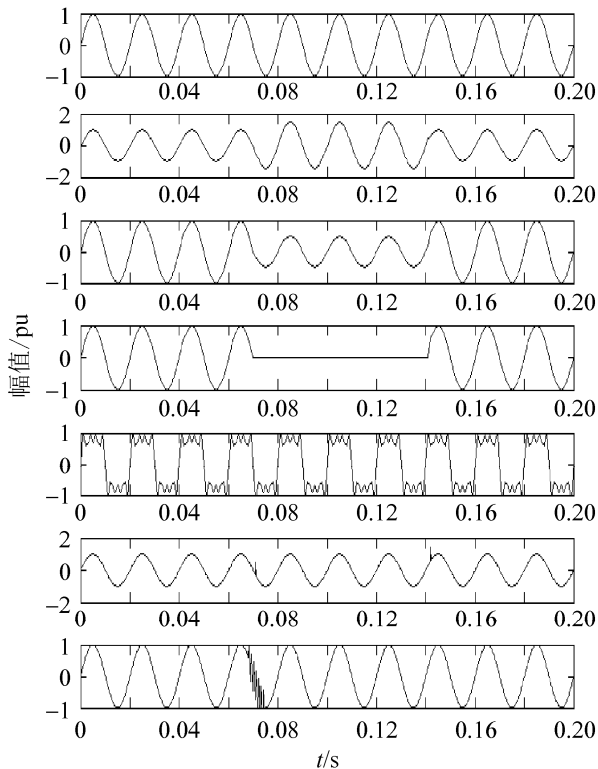


图 1 PQD 波形图

Fig. 1 Disturbance signal waveform

2 基于 MVU 的 PQD 特征提取

2.1 PQD 的小波变换

目前, 采用小波变换提取 PQD 原始信号中的特征向量是一种普遍的方法。多分辨率分析是小波分析的基本理论, Mallat 算法则能快速实现小波变换。其通过重复使用低通滤波器和高通滤波器来实现分解。滤波器得到的低频分量和高频分量各占信号频带的 1/2。对得到的新的低频分量重复上述分解过程, 得到下一层的高频分量和低频分量。

根据 Parseval 定理, 能量小波系数公式如下:

$$\int [f(t)]^2 dt = \sum [a_j(k)]^2 + \sum [d_j(k)]^2 \quad (1)$$

式中: $f(t)$ 为待分解信号; $a_j(k)$ 为小波分解第 j 层的近似系数; $d_j(k)$ 为小波分解第 j 层的细节系数。对 PQD 进行 J 层分解, $f(t)$ 信号的小波变换近似能量分布和细节能量分布分别定义为:

$$E_{a_j} = \sum_n (a_j(n))^2 \quad (2)$$

$$E_{d_j} = \sum_n (d_j(n))^2 \quad (3)$$

式中: $j=1, 2, \dots, J$, 经过 J 层小波分解, 得到 $J+1$ 个特征量, 将特征量构成能量函数 $E=[E_{a_1}, E_{d_1}, E_{d_2}, \dots, E_{d_J}]$ 。

与正常信号相比, 各种 PQD 对应的各尺度下小波变换系数的能量是不同的, 因此可对扰动进行多尺度分解, 然后根据能量特征构造特征向量。由于 Daubechies (db) 系列小波的正交性、紧支性, 对不规则信号较为敏感, 非常适合对 PQD 信号进行变换, 其中根据文献[12-13]可得 db4 小波处理能力最优。用以上提到的 db 正交小波基对采样率为 6.4k Hz 的纯正弦信号(不含噪声)进行分析, 采用 db1-10 小波分别对其进行 10 尺度分解, 如图 2 所示。如图 2 可得在第 7 分解尺度上出现信号分解能量最大能量值。利用以上公式进行 db4 小波 7 层分解, 只提取细节能量部分由 7 个特征量组成向量 $E_D=[E_{d_1}, E_{d_2}, \dots, E_{d_7}]$, 构造 PQD 信号的原始向量集, 作为下一步 MVU 运算的数据样本。

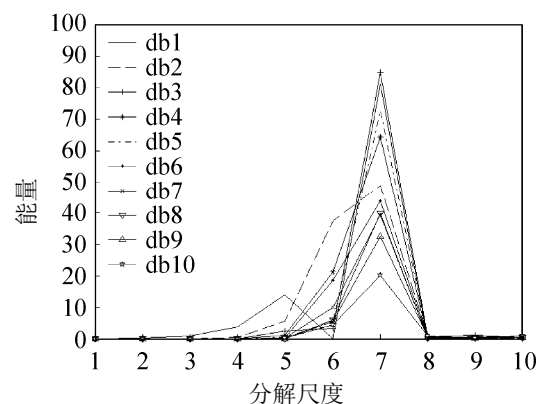


图 2 正弦信号 db1~db10 小波分解的能量分布图

Fig. 2 Energy distribution of sine signal by db1~db10 wavelet decomposition

2.2 MVU 算法

假设高维空间 R^d 中的观测数据集 $X = (x_1, x_2, \dots, x_n)^T$ 是从嵌入在 d 维空间中的 r 维流形 M 上采样得到, 流形学习就是在没有任何关于 M 和 r 先验知识的条件下, 根据高维观测数据集 X 发现未知映射 $f: R^d \rightarrow R^r (r \ll d)$, 并找到与高维观测数据集 X 一一对应的低维流形表示 $Y = (y_1, y_2, \dots, y_n)^T, y \in R^r$ 。

MVU 算法首先根据 X 构造一个 $n \times n$ 的邻接矩阵 W , 如果 x_i 是 x_j 的 k 近邻, 则 $W_{ij}=1$, 否则 $W_{ij}=0$ 。MVU 算法可表述为如下的优化问题:

$$\begin{aligned} \max \quad & \sum_{i,j=1}^n \|y_i - y_j\|^2, \\ \text{s.t.} \quad & \|y_i - y_j\|^2 W_{ij} = \|x_i - x_j\|^2 W_{ij}, \\ & \sum_{i=1}^n y_i = 0. \end{aligned} \quad (4)$$

第一个约束函数是局部等距约束, 它保证高维空间中的数据点向低维映射时, 近邻点间的欧式距离保持不变, 数据集的局部结构因此得以保留。第二个是中心化约束, 用来消除平移自由度。

一个在平方等式约束下的非凸二次规划问题, 容易陷入局部最优解。通过引入核函数可转化为一个凸半正定最优化问题, 定义数据集 Y 的核矩阵为 $K = [K_{ij}]_{n \times n}$, 它的元素为 $K_{ij} = \langle y_i, y_j \rangle$, 其中 $\langle \bullet, \bullet \rangle$ 表示求内积。此时, 上述优化问题可转化为:

$$\begin{aligned} \max \quad & \text{trace}(K) \\ \text{s.t.} \quad & K \geq 0, \\ & \sum_{i,j=1}^n K_{ij} = 0, \\ & K_{ii} - 2K_{ij} + K_{jj} = \|x_i - x_j\|^2, \text{ 如果 } W_{ij} = 1 \end{aligned} \quad (5)$$

式中: $\text{trace}(\bullet)$ 表示矩阵的迹, 第一个增加的约束 $K \geq 0$ 表示 K 为半正定矩阵, 用来保证数据来自于凸集。

对上述的凸半正定最优化问题公式(5)的求解有成熟的快速算法。

设 K^* 为该优化问题的最优解, 对 K^* 进行谱分解得到

$$K_{ij}^* = \sum_{\alpha=1}^n \lambda_{\alpha} V_{\alpha i} V_{\alpha j} \quad (6)$$

式中: λ_{α} 是矩阵 K^* 的第 α 个特征值; $V_{\alpha i}$ 为相应特征向量的第 i 个元素。由此, 高维数据 x_i 的 n 维映射 y_i^* 的第 α 个元素为

$$Y_{\alpha i}^* = \sqrt{\lambda_{\alpha}} V_{\alpha i} \quad (7)$$

至此, d 维观测数据 x_i 的 r 维嵌入可用式(8)表示:

$$x_i \rightarrow y_i = (\sqrt{\lambda_1} V_{1i}, \sqrt{\lambda_2} V_{2i}, \dots, \sqrt{\lambda_r} V_{ri})^T \quad (8)$$

MVU 是一种能够自动学习训练数据核矩阵的流行算法。其通过求解一个半正定规划, 从而得到训练数据的最优核矩阵。在满足 k -最近邻点间局部距离不变的约束下, 使映射到核特征空间中的训练数据方差最大化。因此, 潜在的高维数据中的非线性结构被映射到对应的核特征空间中, 并使得映射为线性。由于 MVU 算法只需要保持数据集局部等距离, 因此 MVU 可以对非凸流形进行有效处理, 且信息量更集中。由于 MVU 算法取核函数时充分考虑了高维数据集的分布特性, 从而保证非线性分布的数据集映射到特征空间后能满足线性分布, 并较好地保持了训练数据集的分布边界。

综上所述, 基于 MVU 的 PQD 信号特征提取算法步骤归纳如下。

- 1) 选择合适的小波函数对原始 PQD 信号进行分解得到 m 维相空间;
- 2) 选择合适的近邻数 k 来进行构建邻接矩阵 $W_{N \times N}$, 若 x_i 是 x_j 的 k 近邻, 则 $W_{ij}=1$, 否则 $W_{ij}=0$, 其中 x_i 为相空间数据点, N 为相空间内数据点个数;
- 3) 求解优化问题公式(5)得到最优核矩阵 K ;
- 4) 对 K 进行谱分解, 根据 K 特征值及其相应特征向量由公式(7)和(8)计算出高维流形在 r 维空间的坐标表示。

2.3 MVU 算法的 PQD 特征提取参数选取

基于 MVU 算法的 PQD 特征提取算法, 根据上一节的介绍可得两个重要参数, 即估计维数参数 r 和邻近点个数 k 。本节分别详细阐述了基于 PQD

信号的本征维数估计和邻域参数选取的方法。

1) 本征维数估计

在确保无信息丢失的前提下,通过最少变量个数表示数据自由变量,称为该数据集的本征维数,其是数据集的一种固有属性。相空间内相点邻域内的分布特质是由信号的本征维数决定的,因此,通过分析相点邻域所张成的子空间可以确定信号的本征维数。在重构相空间中,噪声分布在整个高维相空间中,而有用信号实质上是分布在以本征维数为大小的特征空间里。因此,在采用流形学习进行降维时,若约简目标维数选择过大,则可能消除噪声分布不理想;而若约简目标维数选择过小,则可能消除掉有用信号。

极大似然估计法(Maximum Likelihood Estimate, MLE) 估计数据的本征维数是通过建立近邻点间距离的似然函数估计数据的本征维数。设 $x_1, x_2, \dots, x_n$ 为高维的样本数据,构造一个二项随机过程,建立邻近点距离的似然函数。对于给定的 k , 将本征维数的估计式遍历 $x_1, x_2, \dots, x_n$ 从而得到 n 个局部本征维数的估计值 $m_k(x_i)$, 并取平均值作为全局本征维数的估计值:

$$r = \frac{1}{n} \sum_{i=1}^n m_k(x_i) \quad (9)$$

对每种 PQD 信号的 E_D 进行 MLE 本征维数估计, 结果得到估计维数 $r=3$ 。

2) 邻域参数选取

邻域选择参数, 邻近点个数 k 对于流形学习算法的性能影响也很大。这是因为大量的最近邻域可以促成流形小规模结构的消除及整个流形的平滑。相反, 太少邻域可能误将连续的流形划分成脱节的子流形。由于算法的参数较少且相互独立, 搜索的复杂度不高, 所以在搜索范围内遍历所有的邻近点个数 k , 在一定的约束条件下找到最合适的参数值, 从而搜寻到最优的邻域参数 k 。

对于 PQD 信号的 E_D 进行 MVU 流行算法特征提取计算时, 当 $r=3$ 时, 参数 k 在 10~20 的搜索范围内遍历邻近参数 k , 搜索步长为 2, 其中约束

条件为以 SVM 分类器分类时训练准确率和测试准确率最高, 从而搜寻到最优的邻域参数 k 。参数 k 的寻优结果如图 3 所示。由图 3 中可以得到, k 值在 10~20 之间变化时, 当 $k=12, k=14$ 时, 模型训练准确率和测试准确率最高, 分别为 100%, 99.52%。

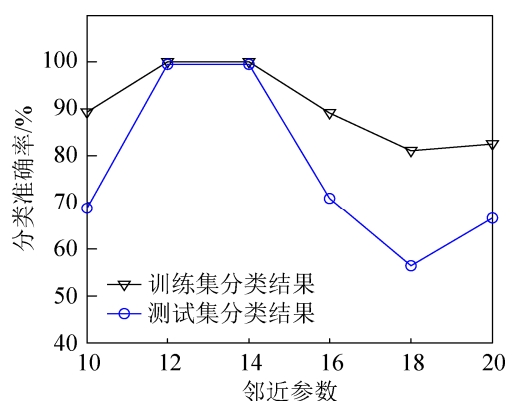


图 3 取不同邻近数 k 值的 PQD 分类结果
Fig. 3 PQD classification results with different k

为了避免计算资源消耗过多, 只要 k 值能至少较好地保持数据集局部低维流形的特性就可以。表 2 对比了当 $k=12, k=14$ 时的 700 个 PQD 信号的特征提取运行时间, 可以看出 $k=12$ 时的耗时要少。因此, 在实验过程中选取 MVU 算法参数为 $r=3, k=12$ 。

表 2 运行时间对比
Tab. 2 Running time comparison

邻近参数	$k=12$	$k=14$
运行时间/s	10.065 720	10.136 792

3 基于 MVU 的 PQD 识别框架

非线性流行学习理论表明, 当原始 PQD 特征集需要压缩数据和分类信号类别时, MVU 流行学习算法可以很好找到潜在在 R^d 内的 R^r 。而 K^* 进行谱分解得到特征值和相对应的特征向量, 可以表征原来的向量空间, 而选取的 $r < d$, 也就是说 r 维特征向量保留了 d 维原始 PQD 信号 X 的所有有用特征, 且信息量更集中。以上分析可以得出, 基于 MVU 算法提取到的特征在分类识别过程中, 要优

于特征分散的原始特征集。因此,不同于传统方法,基于 MVU 算法提取 PQD 特征向量,作为分类器的输入,是一种有改进的 PQD 识别方法,其流程如图 4 所示。

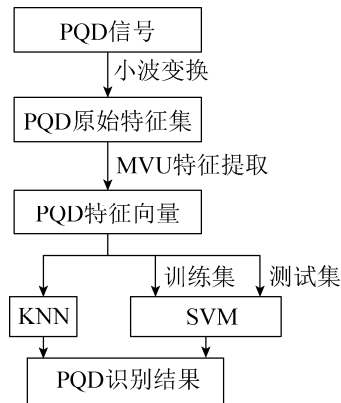


图 4 PQD 识别流程图

Fig. 4 Flow chart of PQD recognition

在特征提取过程中,根据 2.1 节分析可得,小波变换提取原始特征集时选取 db4 小波进行 7 层分解得到 E_D 。根据 2.3 节分析可得, MVU 算法获取 PQD 特征向量时其参数选取为 $r=3$, $k=12$ 。

为了验证本方法特征向量的有效性和鲁棒性,分别选取经典的分类算法 K 最近邻法(KNN, K-Nearest Neighbor)和支持向量机法(Support Vector Machine, SVM)进行分类验证。KNN 分类算法是数据挖掘分类技术中最简单的方法之一。其优点在于其算法简单,容易理解,易于实现,不需要估计参数和分类训练。SVM 是在统计学习理论的基础上发展起来的一种新的模式识别方法。本文选用一对一 SVM 构造多个二分类器实现 PQD 分类,其中选取径向基函数中的高斯核函数作为 SVM 的核函数。由于 SVM 算法需要预学习,因此分别输入训练样本和测试样本进行实验分类。

4 实例验证与分析

4.1 实验数据

根据 1 节中的表 1 所提供的信号模型产生 PQD 数据。取电压正弦信号和电压凸起、电压凹

陷、电压间断、谐波、电压暂态,其中电压暂态包括脉冲暂态和振荡暂态 6 类扰动信号各 100 个,共 700 个样本;输入信号的分析时间长度取 10 个周波(0.2 s),采样率为 6.4 kHz,信号的采样点为 1280 个,每个周波采 128 个点;为较好地模拟实际中的各种情况,保证分析结果的可靠性,令每种扰动的参数如扰动起止时间、幅值、持续时间在允许范围内随机变化,并给扰动信号加入信噪比为 30 dB 的随机白噪声。对 PQD 信号进行采样得到 7 组 100×1280 的样本数据。应用上述数据,实现基于 MVU 算法的 PQD 特征提取,并分别选用 KNN 和 SVM 分类器进行算法验证。在 SVM 分类器分类时,分别取每种扰动 70 个样本作为训练集,取 30 个样本作为测试集。

4.2 PQD 特征提取

为了验证 MVU 方法能有效解决 PQD 特征提取的相关性和冗余性问题,采用 MVU 算法对 E_D 进行数据处理,从而提取 PQD 的 3 维特征向量。并将 ISOMAP, LLE, LTSA, LPP 算法与本文算法进行了特征提取对比。首先,通过 db4 小波 7 层分解的能量值构造样本向量数据,然后再分别采用 Isomap、LLE、LTSA、LPP、MVU 算法进行数据处理,从而提取 PQD 的 3 维特征向量。其可视化嵌入图,分别如图 5~9 所示。其中,不同类别 PQD 的样本表示方法如图例所示。

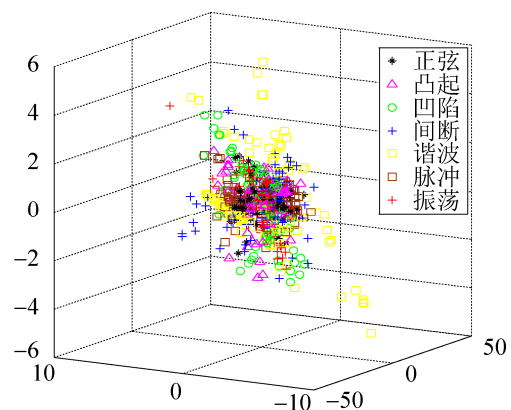


图 5 ISOMAP 嵌入图

Fig. 5 The embedding results by ISOMAP

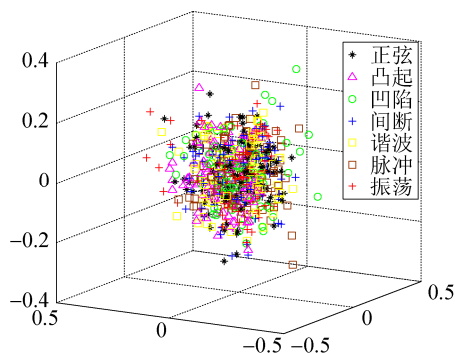


图 6 LLE 嵌入图

Fig. 6 The embedding results by LLE

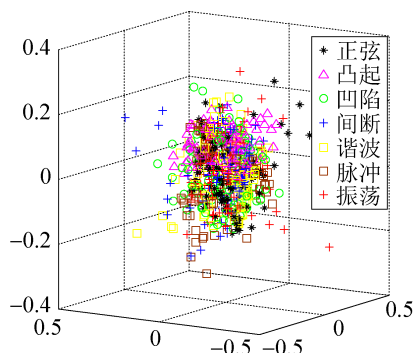


图 7 LTSA 嵌入图

Fig. 7 The embedding results by LTSA

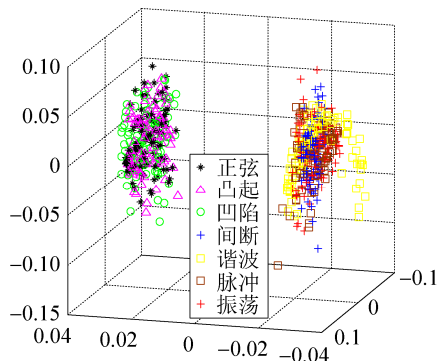


图 8 LPP 嵌入图

Fig. 8 The embedding results by LPP

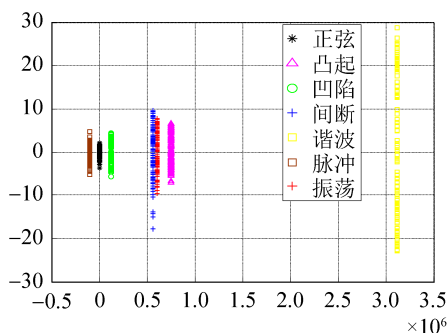


图 9 MVU 嵌入图

Fig. 9 The embedding results by MVU

图 5~7 分别为采用 Isomap, LLE, LTSA 算法进行 PQD 特征提取的可视化嵌入图。可看出以上 3 种降维方法的嵌入结果很不理想, 图中均出现了严重的数据拥挤问题, 嵌入结果中不同类别的样本相互交叠在一起, 几乎无法展示相互分离的状况。

图 8 为采用 LPP 方法进行 PQD 特征提取的三维散点图。该图可以看出 LPP 方法将 PQD 信号分为两类混叠, 即电压正弦、凸起和凹陷信号拥挤混叠在一起, 无法区分; 其他信号拥挤混叠在一起, 无法区分。

图 9 为 MVU 算法的 PQD 特征向量提取结果的嵌入图。由于 MVU 算法更好的保持了训练数据集的分布边界, 三维散点降维结果分离程度较好, 为了更清晰的显示结果可将其三维结果投影到二维平面。由图 9 可得, 7 种不同类别的样本数据降维后的空间分布, 正常信号和 6 种 PQD 信号均能明显区分, 在完成维数简化的同时很好的保持了训练数据的分布边界。

与其他经典的非线性流行学习算法相比, MVU 算法不但完成了对 PQD 特征向量的降维, 而且很好的解决了流行算法在 PQD 分析中出现的数据“拥挤问题”。这是由于 MVU 算法尝试通过学习核矩阵来解决如何选择核函数这个问题。MVU 算法在数据上定义一个邻域图并在结果图中保留两两距离来学习核矩阵。在邻域图中的距离保留不变的约束下(即在数据流行的局部几何形状不变的约束下), 通过使数据点之间的欧几里得距离最大化来实现流行展开。所得到的优化问题可以使用半正定规划来解决。

综上所述, 基于 MVU 算法的 PQD 特征提取方法能保持样本高维空间中隐藏的低维结构, 并将这潜在的流形嵌入到低维空间中。并且 MVU 降维后的低维数据较好的保持了训练数据的分布边界。从以上实验结果分析可以看出 MVU 算法对于 PQD 特征提取的降维效果优于其他流行学习算法。

4.3 PQD 分类

为了验证提出的方法的有效性和鲁棒性, 分别

采用 KNN 和 SVM 分类器, 对以上得到的 PQD 特征向量进行分类验证。

1) KNN

采用 KNN 分类算法来验证基于 MVU 算法的 PQD 特征提取的识别效果。在实验过程中, 在给定范围内取任意 k 值, 基于 MVU 算法得到的 PQD 特征向量的识别结果均为 100%。以上结果说明该特征向量具有明确的分类边界, 使得对于后续的分类算法要求极低, 对于分类器参数选择并不敏感, 是一种强壮稳健的 PQD 特征。

2) SVM

采用一对一方法构造多分类 SVM 分类器, 来验证基于 MVU 的 PQD 特征提取的识别效果。SVM 的核函数采用高斯径向基核函数。其中核函数的尺度参数取 $\sigma=0.1$ 、正则化参数取 $c=100$ 。每种扰动的前 70 个数据用于 SVM 分类器的训练。后 30 个数据用于测试分类的准确率。SVM 分类器的训练集分类结果为 100%。SVM 分类器的测试集分类结果为 99.52%, 如表 3 所示。

表 3 基于 MVU 算法的 SVM 测试集分类结果
Tab. 3 Classification results of SVM test sets based on MVU

PQD	正弦	凸起	凹陷	间断	谐波	脉冲	振荡
正弦	30	0	0	0	0	0	0
凸起	0	30	0	0	0	0	0
凹陷	0	0	29	0	1	0	0
间断	0	0	0	30	0	0	0
谐波	0	0	0	0	30	0	0
脉冲	0	0	0	0	0	30	0
振荡	0	0	0	0	0	0	30

由上节 PQD 特征提取可知, 对 PQD 信号进行采样得到 7 组 100×1280 的样本数据。经过 7 层小波分解, 每种扰动的小波能量数据为 100×7 的矩阵, 通过 MVU 算法降维后得到 100×3 的矩阵。从而达到样本降维的目的。分别选取 MVU 算法降维前 $X_{100 \times 7}$ 为特征向量和 MVU 算法降维后 $Y_{100 \times 3}$ 为特征向量进行 PQD 分类对比, 两种特征向量的识别结果如表 4 所示。

表 4 MVU 算法降维前后 PQD 识别结果

Tab. 4 Comparison of PQD Recognition Results

	维数	训练准确率/%	训练时间/s	测试准确率/%	测试时间/s
降维前	7	71.84	2.398 465	48.10	0.050 456
降维后	3	100.00	0.748 606	99.52	0.010 804

由以上实验结果可得, 在 MVU 算法降维前, 将小波能量 E_D 作为特征向量, 用 SVM 分类器对 PQD 分类, 其训练和测试准确率都低于 MVU 算法, 而其训练和测试运行时间都高于 MVU 算法。实验结果说明由于 MVU 算法处理后的 PQD 特征向量减少了维数, 从而在识别过程中缩减了运行时间; 由于 MVU 算法处理后的 PQD 特征向量很好的保持了训练数据的分布边界, 且信息量更集中, 从而有助于提高分类准确率。

4.4 实测数据验证

通过电力系统中的实际测量数据, 进一步验证 MVU 算法对 PQD 信号识别的有效性。采用国内 500 kV 输电线故障录波数据对 PQD 识别进行验证。所用实测 PQD 数据已经过录波设备采集、处理并记录。每种 PQD 实测数据取 20 组, 选用 SVM 算法进行分类, 其中每种信号的 15 组数据用来训练, 5 组数据用来测试, 其测试分类结果如表 5 所示。

表 5 实测数据分类结果
Tab. 5 Classification result of real data verification

PQD 类型	分类正确率/%
标准	100
凸起	100
凹陷	80
间断	100
谐波	80
脉冲	100
振荡	100

由表 5 可得, 采用本文方法进行实测 PQD 分类仍然可得到较好的分类正确率, 进一步验证了本文所提方法的有效性。

5 结论

本文提出的基于 MVU 流行学习方法的 PQD 特征提取方法, 在考虑了随机参数和噪声的影响下, 较好地完成了对 PQD 信号的低维嵌入表示, 是一种理想的特征提取方法。通过构造通用分类器, 验证了 MVU 算法提取到的 PQD 特征向量的有效性, 其可以减少分类运行时间和提高 PQD 识别率。MVU 维数约简后的低维 PQD 特征向量很好的保持了训练数据的分布边界, 且信息量较其他流行算法更集中, 在降维的同时可以减轻后续 PQD 的分类压力。下一步的研究将尝试该方法在更多复合扰动类型的 PQD 信号识别中的应用。

参考文献:

- [1] 王远, 陶烨, 袁军等. 一种基于 HBase 的智能电网时序大数据处理方法[J]. 系统仿真学报, 2016, 28(3): 559-568.
Wang Yuan, Tao Ye, Yuan Jun, et al. Approach to Process Smart Grid Time-Serial Big Data Based on HBase[J]. Journal of System Simulation, 2016, 28(3): 559-568.
- [2] 宫夏屹, 李伯虎, 柴旭东, 等. 大数据平台技术综述[J]. 系统仿真学报, 2014, 26(3): 489-496.
Gong Xiayi, Li Bohu, Chai Xudong, et al. Survey on Big Data Platform Technology[J]. Journal of System Simulation, 2014, 26(3): 489-496.
- [3] 徐心怡, 贺兴, 艾芊, 等. 基于随机矩阵理论的配电网运行状态相关性分析方法[J]. 电网技术, 2016, 40(3): 5414-5420.
Xu Xinyi, He Xing, Ai Qian, et al. A correlation analysis method for operation status of distribution network based on random matrix theory[J]. Power System Technology, 2016, 40(3): 5414-5420.
- [4] 黄南天, 彭华, 蔡国伟, 等. 电能质量复合扰动特征选择与最优决策树构建[J]. 中国电机工程学报, 2017, 37(3): 776-785.
Huang Nantian, Peng Hua, Cai Guowei, et al. Feature selection and optimal decision tree construction of complex power quality disturbances[J]. Proceedings of the CSEE, 2017, 37(3): 776-785.
- [5] 徐会法, 张明智, 胡晓峰, 等. 基于分数阶傅里叶变换的雷达侦察系统仿真[J]. 系统仿真学报, 2014, 26(5): 1084-1089.
Xu Huifa, Zhang Mingzhi, Hu Xiaofeng, et al. Radar Reconnaissance System Simulation Based on Fractional Fourier Transform[J]. Journal of System Simulation, 2014, 26(5): 1084-1089.
- [6] 李刚, 许鹏程, 韩龙美. 基于小波分析的电力系统故障时空检测与诊断[J]. 系统仿真学报, 2015, 27(12): 3018-3024.
Li Gang, Xu Pengcheng, Han Longmei. Fault Spatial-temporal Detecting and Diagnosis for Power Grid Based on Wavelet Analysis[J]. Journal of System Simulation, 2015, 27(12): 3018-3024.
- [7] Brahma P P, Wu D P, She Y Y. Why Deep Learning Works: A Manifold Disentanglement Perspective[J]. IEEE Transactions on Neural Networks and Learning Systems (S2162-237X), 2016, 27(10): 1997-2008.
- [8] Weinberger K Q, Sha F, Saul L K. Learning a kernel matrix for nonlinear dimensionality reduction[C]// Proceedings of the 21st international conference on Machine learning. USA: ACM, 2004: 106-113.
- [9] 陈如清. 基于改进 MVU 的非线性动态过程故障检测方法[J]. 仪器仪表学报, 2013, 34(9): 2111-2117.
Chen Ruqing. Improved MVU based fault detection method for nonlinear and dynamic process[J]. Scientific Instrument, 2013, 34(9): 2111-2117.
- [10] 张赞, 李本威. 基于最大方差展开的非线性信号降噪方法及其在故障诊断中的应用[J]. 中国科学: 技术科学, 2010, 40(8): 940-945.
Zhang Yun, Li Benwei. Noise reduction method for nonlinear signal based on maximum variance unfolding and its application to fault diagnosis[J]. Science China Technical Science, 2010, 40(8): 940-945.
- [11] Gargoom A M, Ertugrul N, Soong W L. Automatic classification and characterization of power quality events[J]. IEEE Transactions on Power Delivery (S0885-8977), 2008, 23(4): 2417-2425.
- [12] 黄振威, 郑寿森, 祁新梅, 等. 基于 db4 小波变换和 ARM Cortex-M 的电能质量分析系统[J]. 电测与仪表, 2013, 50(1): 83-86.
Huang Zhenwei, Zheng Shousen, Qi Xinmei. Power Quality Analysis Based on db4 Wavelet Transform & Dual ARM Cortex-M[J]. Electrical Measurement & Instrumentation, 2013, 50(1): 83-86.
- [13] 曲正伟, 王婷婷, 郝宏伟, 等. 基于提升 db4 复小波暂态电能质量扰动检测与分类[J]. 电测与仪表, 2014, 51(8): 26-31.
Qu Zhengwei, Wang Tingting, Xi Hongwei, et al. Detection and Classification of Transient Power Quality Disturbance Based on lifting db4 Complex Wavelet[J]. Electrical Measurement & Instrumentation, 2014, 51(8): 26-31.